Plant Breeding **124**, 44—53 (2005)

ISSN 0179-9541

Conformity and genetic relatedness estimation in crop species having a narrow genetic base: the case of cucumber (*Cucumis sativus* L.)*

J. E. Staub[1], S.-M. Chung[2] and G. Fazio[3]

[1] USDA/ARS, Vegetable Crops Unit, Department of Horticulture, 1575 Linden Dr, University of Wisconsin, Madison, WI 53706, E-mail: jestaub@facstaff.wisc.edu; [2] Department of Biochemistry and Cell Biology, State University of New York, Stony Brook, NY 11794; [3] USDA/ARS, Curator, Plant Genetics Research Unit, Geneva, NY 14456, USA

With 4 figures and 3 tables

Received March 15, 2004/Accepted September 22, 2004
Communicated by W. E. Weber

Abstract

A set of 155 SSR (107) and SCAR (48) markers were used to evaluate 53 cucumber (*Cucumis sativus* L.) accessions of diverse origin to characterize genetic relationships and to define a standard marker array that was most effective in detecting genetic differences in this germplasm array. A multivariate marker-based analysis of diverse germplasm using this standard marker array (17 SSR and 5 SCAR markers) was compared with results from a set of 70 previously reported RAPD markers, and then used to explore the potential value of these genetic markers for plant variety protection (PVP) and the establishment of essential derivation (ED) threshold values in this species using elite lines and hybrids and backcross progeny. Diversity analysis allowed identification of distinctly different lines that were used for the construction of three sets of backcross families (BC_1-BC_3). While general genetic relationships among accessions were similar in SSR/SCAR analyses ($r_s = 0.65$) using two genetic distance (GD) estimators, differences in accession relationships were detected between RAPD and SSR/SCAR marker evaluations regardless of the estimator used. The GDs among elite germplasm with known pedigrees were relatively small (0.06–0.23 for any pairwise comparison). GD values decreased and degree of fixation (at three to seven loci depending on the mating) increased with increased backcrossing such that recurrent parent allelic fixation occurred in least one family of each of the BC_3 families. In many instances the degree of fixation of loci was not uniformly achieved in the BC_3. Although the level of genetic polymorphisms will likely restrict the use of molecular markers for PVP and the establishment of ED values, the use of single nucleotide differences will likely provide opportunities to define specific functional distances that have potential for PVP in cucumber. Nevertheless, without an expanded, genetically robust standard marker array (e.g. 50 codominant markers), ED threshold values will be difficult to define in this species, and perhaps will require the appraisal of single nucleotide polymorphisms as discriminators of difference in this species.

Key words: *Cucumis sativus* — essential derivation — functional genetic distance — genetic distance — plant variety protection

Intellectual property rights in the United States can be protected by the Plant Patent Act of 1930, utility patents, and the Plant Variety Protection Act of 1970 as amended in 1994 (Staub et al. 1996a, Staub 1999). In Europe, protection is afforded by Plant Breeders Rights that is dependent upon the Union pour la protection des obtentions vegetales (UPOV 1990,1991) guidelines. Potential infringement of intellectual property related to germplasm necessitated the UPOV to establish the concept of an essentially derived variety (EDV) in 1991 (UPOV 1990, 1991, Smith et al. 1995). The EDV is distinct from the initial variety (IV) from which it was predominantly derived, but conforms in its expression with the essential characteristics of the IV.

A description of genetic differences between the IV and EDV can assist in defining when 'the essential part of the genome of an initial variety has been included in the new variety' (i.e. degree of conformity) (ASSINSEL 1999). Defining 'essential derivation' (ED) requires the development of commonly recognized genetic thresholds that depict the parental and/or ancestral nature of released germplasm in relation to the IV. Therefore, attempts have been made to establish potential crop-specific thresholds for ED of plant varieties.

In Europe, an EDV is judged to be distinct from an IV by national testing authorities (ASSINSEL 1999, Gilliland et al. 2000). Once defined as an EDV, the breeder of the IV is responsible for establishing genetic relationships between the EDV and the IV. There are, however, currently no universally accepted protocols for determining such relationships. Standardization of procedures that work to describe the ancestral nature between an initial and 'new' variety requires estimates of genetic distance (GD) based on morphological and/or molecular criteria (i.e. the distinctness and conformity elements) and, where possible, pedigree analysis (i.e. the predominantly derived element). Appraisals of such relationships for plant variety protection or infringement determinations are enhanced when the genetic relationship of such germplasm is measured against the backdrop of variation within the species itself. This can be accomplished by comparative genetic analyses of an IV and its EDV using a standard germplasm reference array that circumscribes the genetic variation of exotic and adapted germplasm within the species (Staub et al. 1996a, Staub 1999). Such characterizations become more difficult when a species has a narrow genetic base and its inherent polymorphism level is relatively low.

Genetic distance estimation using molecular markers can be used to assist in the determination of ED thresholds by inbred line and varietal relationships (Lombard et al. 2001). Cucumber (*Cucumis sativus* L.; $2n = 2x = 14$) is an economically important processed or fresh market vegetable, but has a

* The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked advertisement solely to indicate this fact. Mention of a trade name, proprietary product, or specific equipment does not constitute a guarantee or warranty by the USDA and does not imply its approval to the exclusion of other products that may be suitable.

narrow genetic base (3–12% polymorphism; Dijkhuizen et al. 1996, Staub and Bacher 1997). Nevertheless, molecular marker analysis has been found useful for describing its genetic diversity for germplasm management (Meglic et al. 1996, Horejsi and Staub 1999) and for variety identification (Dijkhuizen et al. 1996, Meglic and Staub 1996). Bernet et al. (2003) have recently used random amplified polymorphic DNA (RAPD) to assess marker applicability for the protection of cucumber varieties. Single sequence repeat (SSR) (Fazio et al. 2002), and sequence characterized amplified region (SCAR) (Horejsi et al. 1999) markers are now available for a more comprehensive genetic analysis of cucumber. These markers are genetically defined (i.e. mapped), highly reproducible (Fazio et al. 2002), mainly codominant, and evenly distributed along its genomic length (approximately 750 cM, approximately 350 kb; Staub and Meglic 1993, Fazio et al. 2003).

The unique genetic nature of cucumber offers an opportunity to question the relevance of GD estimation in relation to PVP and the establishment of ED thresholds. Thus, experiments were designed in cucumber to: (1) identify a genetically well-defined, stable set of reference markers; (2) define the relative GD among genotypes of diverse origin, and (3) characterize the GD relationships among genotypes derived from elite breeding lines having closely related pedigrees. Data from experiments described herein were then used to address issues of conformity and genetic relatedness for PVP.

Materials and Methods

Plant materials: Identification of genetic distances among diverse accessions (objectives 1 and 2): Previous research has defined accessions that circumscribe the genetic variation in cucumber (1846 to present as described by Meglic and Staub 1996), and were used to form the initial reference accession array for the preliminary analysis conducted herein (Table 1, nos 37–53) (Knerr et al. 1989, Dijkhuizen et al. 1996, Meglic and Staub 1996, Meglic et al. 1996, Horejsi and Staub 1999, Staub et al. 2002). Differences among elite European cucumber germplasm are not remarkable (Dijkhuizen et al. 1996), and thus elite lines exclusively from Rjik Zwann Seed Company (nos 11–20) were employed along with 24 accessions of diverse origin (nos 11–20, 32, 33, 35, 36, 38–45, 47, 53) to determine whether results using different marker types (RAPD vs. SSR/SCAR) and GD estimators were similar (objective 1) and to assist in defining genetic relationships between European Glasshouse and exotic cucumber accessions (objective 2). Based on this and previous assessments (morphological and molecular; Dijkhuizen et al. 1996, Meglic and Staub 1996, Staub and Bacher 1997, Bernet et al. 2003), an array of modern commercial germplasm (use or release date >1990) spanning diverse market classes was selected for evaluation (Table 1, nos 1–36). This standard germplasm reference array allowed for a more expanded assessment of genetic diversity (objective 2).

Characterization of genetic distances among elite accessions and derived populations (objective 3): To determine genetic relationships between closely related genotypes and populations, specific comparisons were made between commercial elite lines (Table 1, nos 8–10, 22–24, 26, 27) and between a set of three backcross populations derived from US processing cucumber lines. The ancestry of the commercial lines and hybrids used in specific comparisons are proprietary and thus their specific pedigrees were not known to this study. However, some general information regarding some of the accessions used was obtained from commercial seed companies for comparative analysis. Mediterranean type hybrids 8 (EZ-8) and 9 (EZ-9) and 10 (EZ-10) originating from Enza Zaden Seed Company contain a common parent, and the female parents of hybrids 8 and 9 are genetically very similar (i.e. by pedigree). The parents of the Zeraim Gedera hybrid ZG-M (no. 27) are ZG-K (no. 26) and ZG-F (no. 23). Likewise, the parents of hybrid ZG-G (no. 24) are ZG-D (no. 22) and ZG-F (no. 23).

Backcross populations were developed by mating lines WI 5551 (no. 30), G421 (no. 31), and H-19 (no. 29) to line Gy-14 (no. 33) (Table 1). Line Gy-14 was then used as a recurrent parent in F_1 matings to produce the BC_1, and subsequently BC_2, and BC_3 populations. A total of three, four and five backcross families were constructed as a result of initial WI 5551 $\times$ Gy-14, G421 $\times$ Gy-14, and H-19 $\times$ Gy-14 matings, respectively. These parents were chosen based on the results of the diverse accession analysis and/or their use in commercial practice (i.e. objective 2).

Molecular marker and DNA sequencing analysis

Molecular marker analyses: Young leaf tissue (at the two- to three-leaf stage) from each accession was harvested from a sample of at least 15 plants, and then bulked for DNA analysis. DNA was extracted from leaf tissue using a CTAB extraction procedure (Maniatis et al. 1982, Staub et al. 1996b).

Initially, 48 SCARs (Serquen et al. 1997, Horejsi et al. 1999) and 107 SSR (Fazio et al. 2002, 2003) markers were evaluated. After this initial accession screening, the most polymorphic markers (i.e. most variable loci on a per band basis) were designated for inclusion in the standard marker array (objective 1) and used in further genetic analyses (Table 2; objectives 2 and 3). The potential efficacy of this array was then determined by comparison to a set of 70 previously described RAPD markers (Horejsi and Staub 1999) for its ability to resolve genetic relationships among 24 diverse cucumber accessions (see above section).

The number of plants included in each bulk allow for the identification of a heterozygote in a bulk sample with a 95% probability level. Thus, although actual frequencies could not be estimated because DNA was isolated in bulk, values (e.g. GD and fixation) derived from analysis do provide information of heterogeneity at loci. The markers employed are unlinked (Fazio et al. 2003; Table 2), and because of past observations with another codominant marker system (isozymes) using this germplasm (Knerr et al. 1989), segregation of SSR loci in F_2 populations (Fazio et al. 2003) were assumed to be segregating in bulk samples examined at an allelic frequency of $P = 0.5$ and $q = 0.5$. Thus, the determination of frequencies for the estimation of the rates of allelic fixation at polymorphic marker loci was performed simply as the presence or absence of a heterozygote for each bulk according to Widrlechner et al. (1992).

The RAPD PCR and electrophoresis were performed according to Horejsi et al. (1999). Likewise, optimized PCR, electrophoresis, staining, and analysis by SCAR and SSR markers were according to Horejsi et al. (1999) and Fazio (2001) and Fazio et al. (2002), respectively.

DNA sequence analysis: To provide data on the potential value of DNA sequence polymorphisms, an RAPD amplicon associated with multiple lateral branching in cucumber (Serquen et al. 1997) was converted to an SCAR (Horejsi et al. 1999) and designated SCAR $L18_{600}$ (subscript defining is mobility after agarose gel electrophoresis). This marker was monomorphic between the mapping parents H-19 and G421. The SCAR $L18_{600}$ SCAR products were sequenced following the methodology of Fazio (2001) using an ABI 3700 sequencer (Applied Biosystems, Foster City, CA, USA) to determine if the amplicon housed single nucleotide polymorphisms (SNPs).

Statistical analysis: Initial multivariate analyses [multidimensional scaling (MDS), and principal component analysis (PCA)] using a 155-marker (48 SCAR and 107 SSR) data set derived from an evaluation of 53 cucumber accessions (Table 1) were conducted to determine genetic relationships. Jaccard (1908) similarity coefficients were calculated from RAPD and SSR/SCAR data. The Jaccard's

Table 1: Description of accessions used in the genetic analysis of cucumber (*Cucumis sativus* L.) germplasm

No.	Identification and/or name	Hybrid/line/PI[1]	Type[2]	Country of origin[3]	Source[4]	Objective[5]
1	EZ-1	H	EL	The Netherlands	Enza Zaden	2
2	EZ-2	H	EL	The Netherlands	Enza Zaden	2
3	EZ-3	H	EL	The Netherlands	Enza Zaden	2
4	EZ-4	H	EL	The Netherlands	Enza Zaden	2
5	EZ-5	H	EP	The Netherlands	Enza Zaden	2
6	EZ-6	H	EP	The Netherlands	Enza Zaden	2
7	EZ-7	H	EP	The Netherlands	Enza Zaden	2
8	EZ-8	H	MT	The Netherlands	Enza Zaden	3
9	EZ-9	H	MT	The Netherlands	Enza Zaden	3
10	EZ-10	H	MT	The Netherlands	Enza Zaden	3
11	RZ1	L	EL	The Netherlands	Rijk Zwaan	1&2
12	RZ2	L	EL	The Netherlands	Rijk Zwaan	1&2
13	RZ3	L	EL	The Netherlands	Rijk Zwaan	1&2
14	RZ4	L	EL	The Netherlands	Rijk Zwaan	1&2
15	RZ5	H	EL	The Netherlands	Rijk Zwaan	1&2
16	RZ6	L	EL	The Netherlands	Rijk Zwaan	1&2
17	RZ7	L	EL	The Netherlands	Rijk Zwaan	1&2
18	RZ8	L	EL	The Netherlands	Rijk Zwaan	1&2
19	RZ9	L	EL	The Netherlands	Rijk Zwaan	1&2
20	RZ10	L	EL	The Netherlands	Rijk Zwaan	1&2
21	ZG-A	H	NC	Israel	Zeraim Gedera	2
22	ZG-D	H	NC	Israel	Zeraim Gedera	3
23	ZG-F	L	MT	Israel	Zeraim Gedera	3
24	ZG-G	H	WT	Israel	Zeraim Gedera	3
25	ZG-J	H	EP	Israel	Zeraim Gedera	2
26	ZG-K	L	ES	Israel	Zeraim Gedera	3
27	ZG-M	H	ES	Israel	Zeraim Gedera	3
28	ZG-P	H	EP	Israel	Zeraim Gedera	2
29	H-19	L	AP	USA	USDA	2&3
30	WI 5551	L	AP	USA	USDA	2&3
31	G421	L	AP	USA	UW	2&3
32	Chicago Pickling (Chi Pick)	OP	AP	USA	Sand Hill	1&2
33	GY14	L	AP	USA	UW	2&3
34	Poinsett 76	L	AS	USA	CU	2
35	Dasher II	H	AS	USA	Petoseed	2
36	WI 2757	L	ES	USA	UW	1&2
37	Ames 20089	PI	NC	Egypt	NCRPIS	2
38	PI 177359	PI	EP	Turkey	NCRPIS	1&2
39	PI 183967, *C. hardwickii*	PI	WT	India	UW	1&2
40	PI 188749, Baladi	PI	ES	Egypt	NCRPIS	1&2
41	PI 255936, Nidin	PI	EL	The Netherlands	NCRPIS	1&2
42	PI 257486, Yi-Sang	PI	AP	China	NCRPIS	1&2
43	PI 285606, Monastyrski/Free	PI	AP	Poland	NCRPIS	1&2
44	PI 285607, Monastyrski/Ulri	PI	AP	Poland	NCRPIS	1&2
45	PI 369717	PI	NC	Poland	NCRPIS	1&2
46	PI 385967	PI	WT	Kenya	NCRPIS	2
47	PI 432851, 085 F1	PI	EL	China	NCRPIS	1&2
48	PI 451976, Yomaki	PI	NC	Japan	NCRPIS	1&2
49	PI 525153	PI	MT	Egypt	NCRPIS	2
50	PI 525157	PI	ES	Egypt	NCRPIS	2
51	PI 525159	PI	NC	Egypt	NCRPIS	2
52	PI 605927	PI	WT	India	NCRPIS	2
53	PI 606057	PI	WT	India	NCRPIS	2

[1] H, hybrid (commercial or experimental hybrid); L, line (inbred line); OP, open- pollinated variety or maintained by open pollination; PI, plant introduction [cultigen in the US Plant Germplasm System, at North Central Regional Plant Introduction Station, Ames, IA (NCRPIS)].
[2] 'EL' is an European long (greenhouse) type, 'ES' is an European short (Mediterranean) type, 'EP' is a European processing (pickling) type, 'AS' is an American slicing type, 'AP' is an American processing (pickling) type, 'MT' is a European 'Mini' type, 'WT' is a late-flowering wild type, 'NC' is a type that does not fit in any of the categories.
[3] Country from which the seed was received or, in the case of plant introductions, the country where the germplasm was collected.
[4] The commercial seed company (Enza Zaden, Enkhuizen, The Netherlands; Rijk Zwaan, De Lier, The Netherlands; Zeraim Gedera, Gedera, Israel), public institution (UW, University of Wisconsin-Madison, WI; CU, Cornell University, Ithaca, NY, and NCRPIS) which supplied the seed of each accession.
[5] Objective 1 = identify a genetically well-defined, stable set of reference markers; objective 2 = define the relative GD among genotypes of diverse origin, and objective 3 = characterize the GD relationships among genotypes derived from elite breeding lines having closely related pedigrees.

estimator was chosen for marker comparisons based on its previous use in cucumber diversity analyses, its concordance with other GD estimators (Horejsi and Staub 1999, Mliki et al. 2003), and its accepted use in binary data analysis (presence/absence) (Janowitz 1980, Hubalek 1982). The conversion of data into individual pairwise GD estimates among the individuals examined was accomplished by calculating the complement of each coefficient ($1 - J_{ij}$) as described by Spooner et al. (1996).

Table 2: Primer sequences, motif, optimal annealing temperature, and linkage group (trait) associated with simple sequence repeats (SSR) and sequence characterized amplified region (SCAR) markers used to characterize genetic relationships in cucumber (*Cucumis sativus* L.)

Marker type	Linkage group (trait)[1]	Primer 5′ to 3′	Motif	Optimal annealing temperature (°C)
CSWCT02B		F: TTCTGCATACCCTCTCCT R: CACACTTCCAGATGGTTG	(CT)21(TG)8	55–58
CSWCT16B	2	F: CTTATGGTCGGAGAAG R: CTCAGATAACCCAAAATA	(CT)14	58–62
CSWCT25[2]	**1**	**F: AAAGAAATTAAGTCAATCAAACCG R: CCCACCAATAGTAAAATTATACAT**	**(CT)5CTT(CT3)**	**45–63**
CSWCT28	**1 (EAR)**	**F: GAATTCAAAAGCATTTCAAAACTA R: GAATTCAATTGGGTTTTTGAACCC**	**(CT)10(TA)9**	**45–65**
CSWCT30		F: CATGAATCTCAAGTCTTAAACCC R: AAAGGATTGAGAAAGAAATTAAGG	(CT)8	55–65
CSWCTT08		F: GATATAAGCGTTGTGAGGATATGC R: CGTGCTCTATGAAGTAAATTAGTA	(TAA)3(CTT)6	55–65
CSWGAAA02		**F: AGGGCGTGTGAAAATTTGATATAA R: TTCGAGAGTGGAGGGCACTTTCGT**	**(CTTT)8**	**45–65**
CSWGAAT01		F: GTCGGCTTGTGAAGAGAGATTGTG R: GTGGGCACTGGTCAGGCGTTGAGA	(GAAT)5	45–65
CSWGATT01C		**F: TATTGAAACAGAAATTAACATTGG R: TCTTATCCACATTCCATTAAGAAG**	**(GAA)10**	**55–68**
CSWGCA01	1	F: AGTGATGGTGCAGGGCTATCTTAT R: TTGTCTTCCCTCCTCTTCTCGTCT	(GCA)8	55–65
CSWTA05		F: GCATGAGCTCGAGCTGGTGTAGTG R: CGCCTGTTTTCATTTTGATTGGTT	(TA)12	55–65
CSWTA08B	**6**	**F: TTGCATTAATGCTATAAACTTACC R: GAAATTAATATTTAGGCATTG**	**(T)7(TA)7 IN-DEL**	**54–56**
CSWTA09		F: CTACAAAACCTCTCATTCCTTATT R: TCTACTTTTAAATTTAGCACAACT	(TA)4(TG)3(TA)6	55–65
CSWTAAA01	**4 (MLB)**	**F: CAATGCCTCAATCTGATAGGAATG R: ACTGGCTCTCTACATATTGTGAGG**	**(TAAA)4**	**57–65**
AJ18SCAR	4 (MLB)	F: GGCTAGGTGGTATGGGGATGACAT R: GGCTAGGTGGGCTTAAGTTCTTTC	SCAR	50–54
AW14SCAR	**3**	**F: GGTTCTGCTCTTCATTCATTTTCA R: GGTTCTGCTCTAAATAACCAAAAA**	**SCAR**	**56–64**
BC523SCAR	**1**	**F: ACAGGCAGACCCGACGAGGGGCAG R: ACAGGCAGACAAGAGTTTGAGGAT**	**SCAR**	**70**
CS-L18-3SCAR	**1**	**F: CTCTTTCAATCATCTTTCTTCTCT R: ATCATAACAATGATATATTTTACG**	**SCAR**	**45–57**
J5SCAR	**1**	**F: CTCCATGGGGTGCACGTTAACGTT R: CTCCATGGGGCAGCTAAACAGCGG**	**SCAR**	**54–63**
CMGA165	**3**	**F: CTTGTTTCGAGACTATGGTG R: TTCAACTACAGCAAGGTCAGC**	**(GA)10**	**50–51**
NR2	**2**	**F: CTGAAAGCAGTTTGTGTCGA R: AAAGAAGGAAGAGGCTGAGA**	**(CT)12**	**50**
NR60	6	F: AAGCACTTAAATGAGAATCG R: AATAGTAGCCTGTTATATCC	(TG)8(AG)8	46

[1] Linkage group numbers and marker-trait associations (MLB, multiple lateral branching and EAR, relative earliness) according to Fazio et al. 2003.
[2] Bold indicates markers used in cluster analyses of parents and backcross progeny used for relationship determinations (Fig. 3).

To investigate differences in genetic relationships assessed by codominant markers (Jackson et al. 1989), SSR/SCAR data were also subjected to distance estimation according to Nei (1973). Thus, both distance estimators were used in calculating GD among accessions from SSR/SCAR data. Based on this analysis, a distance estimator (i.e. Jaccard's coefficient) was chosen to estimate GD in subsequent relationship analyses (i.e. lines, hybrids, and BC populations). To determine genetic relationships among families, cluster analysis was performed using SSR/SCAR markers based on the number of polymorphic loci detected between BC parents (Sorensen 1948; objective 3). Deviations from predicted progress toward a recurrent parent during backcrossing (e.g. 0.75 at BC_1) were examined using Student's *t*-test.

Unweighted pair-group method using an arithmetic average (UPGMA) cluster analyses were performed on GD matrices in the initial analysis (objective 1), and relationships among accessions were visualized as dendrograms using the NTSYS-pc program version 2.02 (Rohlf 1997). Multidimensional scaling was employed in an expanded diversity analysis (Kruskal and Wish 1978; objective 2), and Spearman's coefficient of rank correlations (r_s; Steel and Torrie 1980) were calculated between genetic difference estimates obtained from RAPD-Jaccard and SSR/SCAR-Jaccard data sets.

Results

Genetic distances among genotypes of diverse origin

The initial analysis of germplasm provided an explanation of a substantial portion of the observed variation according to PCA (66.4% of variation explained by the first three principle components; data not presented) and MDS (stress value = 0.44; Fig. 1). Although both procedures gave similar results, MDS was chosen for depiction of genetic relationships because successive characteristic fixed formula iterations and smoothing by regression analysis provided a clear depiction of accession relationships.

A large group of accessions (35) showed considerable genetic affinities. This group included Turkish (PI 177359), Polish (PI 285607, PI 369717), Chinese (PI 432851, PI 257486), Egyptian

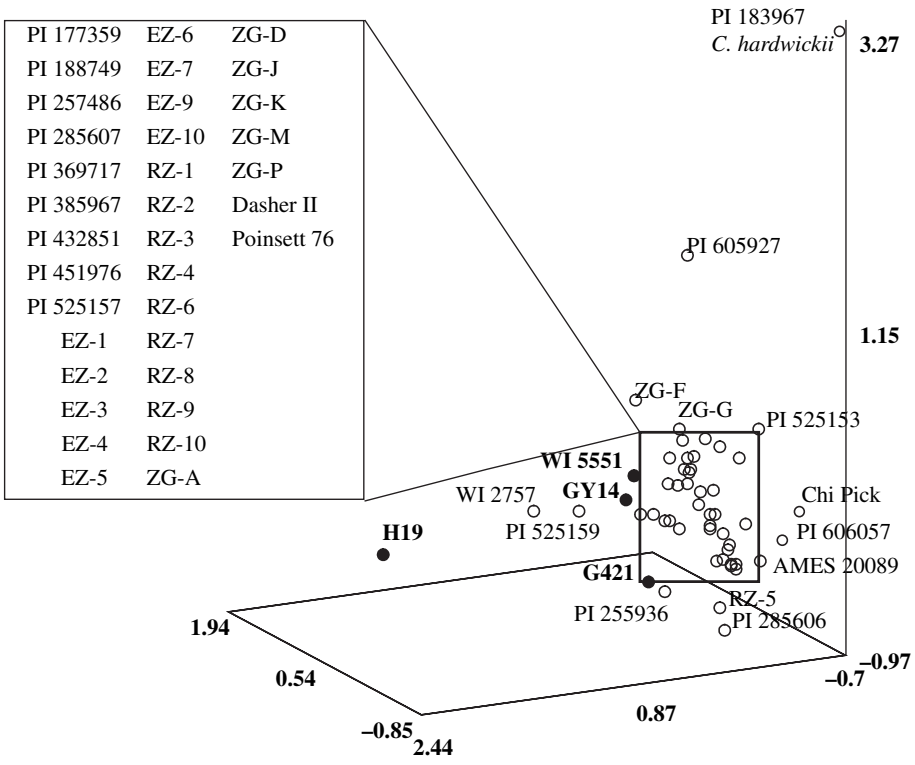

Fig. 1: Depictions of the genetic relationships among 53 *Cucumis sativus* L. accessions using multidimensional scaling of genetic distances as estimated by 107 SSR and 48 SCAR loci as framing criteria (see Table 1 for accession descriptions)

(PI 188749, PI 525157), Kenyan (PI 385967), Japanese (PI 451976), and many elite accessions originating from The Netherlands, Israel, and the United States (i.e., 'Dasher II', 'Poinsett 76'). Unique and peripheral to this group were the accessions PI 183967 (India), PI 605927 (India), PI 606057 (India), RZ-5 (The Netherlands), ZG-F (The Netherlands), PI 255936 (The Netherlands), PI 285606 (Poland), PI 525159 (Egypt), PI 525153 (Egypt), Ames 20089 (Egypt), Chicago Pickling (USA), WI 5551 (USA), GY-14 (USA), WI 2757 (USA), G421 (USA), and H-19 (USA).

Identification of a stable set of reference markers

Of the initial array of 155 SSR/SCAR markers used to evaluate a selected set of 53 accessions, 22 (14%) markers were identified for their inclusion in a standard marker array (Table 2). These markers (17 SSR and five SCAR) were chosen because they were discriminatory among at least five accessions and defined single loci with two alleles (data not presented).

Comparisons between RAPD and SSR/SCAR markers

Average distances between each accession and all others studied indicates differences in GD values depending on marker type and method of distance calculation (Fig. 2). Average distance estimates between any two distance estimators (i.e., RAPD vs. Jaccard, SSR/SCAR-Nei, SSR/SCAR-Jaccard) are tabularized in the figure for each accession examined For example, for RZ-1 the average distance between accessions taken collectively was 0.53 when analyzed by RAPD markers. While the average GD for RAPD-Jaccard and SSR/SCAR-Jaccard was 0.44 and 0.55, respectively, GD based on Nei SSR/SCAR estimation was 0.37. The r_s correlation values between RAPD and SSR/SCAR-Jaccard estimation was 0.65 (data not presented). Albeit correlations cannot be made between values derived from Jaccard and Nei matrices, visual inspection of dendrograms of accession relationships suggest similarities between SSR/SCAR-Nei and SSR/SCAR-Jaccard analyses. This is not the case when RAPD-Jaccard and SSR/SCAR-Nei or RAPD-Jaccard and SSR/SCAR-Jaccard dendrograms are compared. Thus, while general genetic relationships among accessions were similar in SSR/SCAR analyses regardless of the estimator used, differences in accession relationships were detected between RAPD and SSR/SCAR marker evaluations. For instance, in RAPD analysis, elite germplasm was grouped into one branch, while in the SSR/SCAR evaluation the accessions RZ-7 through RZ-10 were dispersed in several branches. The relationships among PIs was relatively consistent regardless of the marker used and mode of GD estimation.

Genetic distance relationships among elite genotypes and their derivatives

The GD (SSR/SCAR-Nei) between the Mediterranean type hybrids no. 8 (EZ-8) and no. 9 (EZ-9), no. 8 and no. 10 (EZ-10), and 9 and 10 was 0.06, 0.13, and 0.06, respectively, confirming their relatively close relationship based on pedigree. In contrast, the GD between no. 26 (ZG-K) and no. 23 (ZG-F), no. 26 and no. 27 (ZG-M), and no. 23 and no. 27 was 0.31, 0.15, and 0.31, respectively. Suggesting larger differences among closely related lines that might have been predicted by pedigree. Likewise, the GD between no. 22 (ZG-D) and no. 23 (ZG-D), no. 22 and no. 24 (ZG-G), and no. 23 and no. 24, was 0.26, 0.07, and 0.23, respectively.

Genetic distance relationships among backcross populations

Based on the analysis of 12 polymorphic SSR/SCAR markers drawn from the standard marker array after an initial analysis

Average genetic distance estimates

Line or hybrid	RAPD-Jaccard	SSR/Scar-Nei	SSR/Scar-jaccard
RZ-1	0.53	0.33	0.59
RZ-2	0.46	0.26	0.65
RZ-3	0.51	0.32	0.60
RZ-4	0.46	0.32	0.60
RZ-5	0.47	0.34	0.58
RZ-6	0.45	0.30	0.61
RZ-7	0.50	0.33	0.58
RZ-8	0.50	0.33	0.58
RZ-9	0.49	0.36	0.55
RZ-10	0.43	0.42	0.50
PI 188749	0.48	0.39	0.52
PI 255936	0.48	0.36	0.57
Chi Pick	0.49	0.38	0.54
Dasher II	0.34	0.43	0.49
WI 2757	0.27	0.46	0.47
PI 183967	0.28	0.83	0.28
PI 285606	0.46	0.46	0.48
PI 432851	0.37	0.30	0.60
PI 257486	0.40	0.39	0.52
PI 285607	0.41	0.34	0.57
PI 369717	0.47	0.31	0.59
PI 451976	0.39	0.38	0.54
PI 177359	0.48	0.28	0.62
GY14	0.44	0.38	0.53
Mean	0.44	0.37	0.55

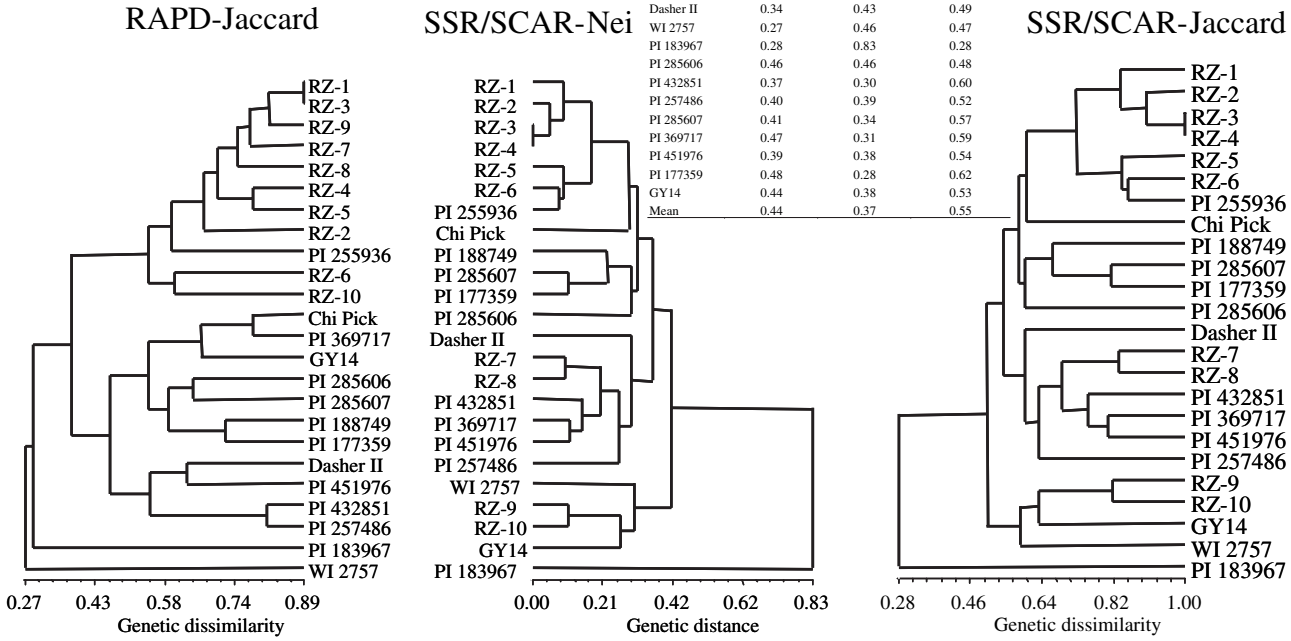

Fig. 2: Cluster analysis (by UPGMA) and average genetic distance among cucumber (*Cucumis sativus* L.) germplasm (see Table 1 for accession descriptions) using Jaccard's (1908) and Nei's (1973) estimation employing a set of 17 SSR/five SCAR and 70 RAPD marker loci

of diverse accessions (Fig. 1), backcross generations were partitioned in five groups (A–E) after cluster analysis (nodes 1–4). While WI 5551 and H-19 possessed genetic distinction (groups D and E, respectively) from the remaining accessions examined, G421 and Gy-14 were placed into group A and B, respectively. Group A was composed exclusively of G421 × Gy-14 progeny families (F_1, BC_1, BC_2, and BC_3). The composition of group B was diverse, consisting of two G421 BC_3 families, a H-19 BC_2 family, three WI 5551 BC_2 families and three BC_3 families, WI 5551 × Gy-14 F_1, the WI 5551 BC_1 family, and three H-19 BC_3 families. Group C was composed exclusively of H-19 × Gy-14-derived families.

Genetic distances among the parents of each backcross family series differed depending on the heterozygosity at the loci examined (Table 3). The GD between WI 5551 and Gy-14, G421 and Gy-14, and H-19 and Gy-14 was 0.24 (at three loci), 0.28 (four loci), and 1.0 (at seven loci), respectively (Table 3). Predictably GD values decreased and degree of fixation usually increased with increased backcrossing such that recurrent parent allelic fixation occurred in least one of each of the BC_3 families. The GD among backcross families within a mating pair reflected the degree of allelic fixation. While a broad range of allelic fixation (0.0–0.71) occurred in H-19 × Gy-14-derived families, the degree of fixation of BC_3 families derived from WI 5551 and G421 ranged between 0 and 1. However, in only four cases [BC_3K (5551 × Gy14), BC_3B (G421 × Gy14), BC_3D (G421 × Gy14), and BC_3H (Gy14 × H19)] did progress toward the recurrent parent meet expectations ($P < 0.05$).

Sequence analysis of lines H-19 and G421

There were 15 SNPs detected during sequence comparisons of amplicons of the $L18_{600}$ SCAR primer using H-19 and G421 as template DNA (Fig. 4).

Discussion

Article 1 of the 1991 UPOV Convention provides a clear concept of variety distinctiveness as being 'defined by the expression of the characteristics resulting from a given genotype or genotypes' (UPOV 1991). It indicates that not only the recognizable expressed parts of the genome (e.g., characteristics of DNA), but also other heritable but indirectly expressed parts of the genome may also be considered to contribute to a variety's definition and essential identity. Technologies such as molecular analysis can be valuable for variety identity analyses when genotypes exist possessing similar genetic pedigrees in crop species that have a genetically narrow germplasm base (Gilliland et al. 2000, de Riek 2001). In cucumber, however, only distantly related germplasm could be unequivocally differentiated by marker analysis (Figs 1–3; e.g. H-19 and G421). Moreover, although the standard marker array was somewhat useful in describing differences between

Table 3: Genetic distance estimates (Nei) and fixation of line GY14 marker alleles among cucumber (*Cucumis sativus* L.) backcross progeny

	5551	GY14	F1 (5551xGY14)	BC1 (5551xGY14)	BC2J (5551xGY14)	BC2K (5551xGY14)	BC2L (5551xGY14)	BC3J (5551xGY14)	BC3K (5551xGY14)	BC3L (5551xGY14)
Group designation[1]	D	B	B	B	B	B	B	B	B	B
Genetic distance to GY14	0.24	0.00	0.12	0.12	0.08	0.08	0.08	0.08	0.00	0.04
Fixation of GY-14 alleles at three loci			0.00	0.00	0.33	0.33	0.33	0.33	1.00	0.67

	G421	GY14	F1 (G421xGY14)	BC1 (G421xGY14)	BC2A (G421xGY14)	BC2B (G421xGY14)	BC2C (G421xGY14)	BC2D (G421xGY14)	BC3A (G421xGY14)	BC3B (G421xGY14)	BC3C (G421xGY14)	BC3D (G421xGY14)
Group designation	A	B	A	A	A	A	A	A	A	B	A	B
Genetic distance to GY14	0.28	0.00	0.16	0.16	0.16	0.16	0.16	0.12	0.08	0.00	0.12	0.00
Fixation of GY-14 alleles at four loci			0.00	0.00	0.00	0.00	0.00	0.25	0.50	1.00	0.25	1.00

	H19	GY14	F1 (GY14xH19)	BC1 (GY14xH19)	BC2E (GY14xH19)	BC2F (GY14xH19)	BC2G (GY14xH19)	BC2H (GY14xH19)	BC2I (GY14xH19)	BC3E (GY14xH19)	BC3F (GY14xH19)	BC3G (GY14xH19)	BC3H (GY14xH19)	BC3I (GY14xH19)
Group designation	E	B	C	C	B	C	C	C	C	B	C	B	B	C
Genetic distance to GY14	1.00	0.00	0.25	0.25	0.16	0.12	0.19	0.16	0.22	0.12	0.08	0.12	0.00	0.12
Fixation of GY-14 alleles at seven loci			0.00	0.00	0.43	0.57	0.29	0.43	0.14	0.57	0.71	0.57	1.00	0.57

[1]According to Fig. 3.

elite parental lines and their F_1 cross progeny (Fig. 2), the associated GDs were predictably small (GD = 0.06–0.31 for any pairwise comparison) and would likely not be of legal consequence where infringement is being investigated. For instance, although codominant markers (i.e., isozymes and RFLP) provide similar ($R^2 = 0.77$) genetic identity information (Dijkhuizen et al. 1996), a standard marker array consisting of 22 isozyme markers was effective in explaining only about 75% of the variation present in a diverse set of cucumber germplasm (Staub et al. 1997). These studies and the data presented herein suggest that molecular markers will continue to be effective for the analysis of genetic diversity in cucumber germplasm collections (Staub et al. 2002) and varietal description for PVP (Staub 1999). However, molecular markers likely will not be adequate, by themselves, for plant patenting or clarification in cases of infringement in cucumber.

The protection of a variety (PVP) often requires a rigorous declaration of uniqueness, and in the case of species with a narrow genetic base such as cucumber, application and cost of development is often an important consideration if markers are used for varietal description. Theoretically, an array of 150 dominant markers could explain about 90% of the variation inherent among diverse cucumber populations (Staub et al. 1997). However, because of technical problems inherent in some marker systems (e.g., RAPD) such markers are not recommended for PVP or the estimation of ED values (Staub et al. 1996b). The disparity in marker system results (i.e., RAPD and SSR/SCAR estimation of GD) detected herein during the initial analysis (Fig. 2) are likely due to the differences in numbers of markers employed in each system. The 22 SSR/SCAR standard marker array use in our study was chosen because of its discriminatory power from a pool of approximately 140 currently available SSR/SCAR markers (Horejsi et al. 1999, Fazio et al. 2002). The development of a highly discriminatory standard marker array consisting of an expanded array of codominant SSR markers (perhaps 50; theoretically explaining about 85–90% of the variation in a diverse germplasm array) is possible and would likely be valuable for PVP of cucumber as its genetic diversity (elite and exotic germplasm) has been rigorously defined (Meglic and Staub 1996, Meglic et al. 1996, Horejsi and Staub 1999). However, the development of such discriminatory marker systems is currently extremely costly (Staub et al. 1997, Fazio et al. 2002).

Marker-based GD differences among elite cucumber cultivars of various market classes are extremely small (Dijkhuizen et al. 1996, Meglic and Staub 1996, Horejsi and Staub 1999; Figs 2 and 3, Table 3). The predictably small GDs detected between the closely related genotypes (i.e., elite commercial parents and F_1 progeny) examined herein typify genetic differences between elite cucumber germplasm and suggest that ED threshold values will be difficult to quantify. Additionally, although the standard marker array used achieved discrimination among BC families to some degree (Fig. 3), discrimination between backcross generations within families was far less effective, and in several instances the predicted degree of fixation of the unlinked loci employed was not uniformly achieved in the BC_3 (Table 3). This lack of fixation was likely due, at least in part, to the influence of sampling, and suggests that perhaps larger population sizes and/or individual plant sampling within a

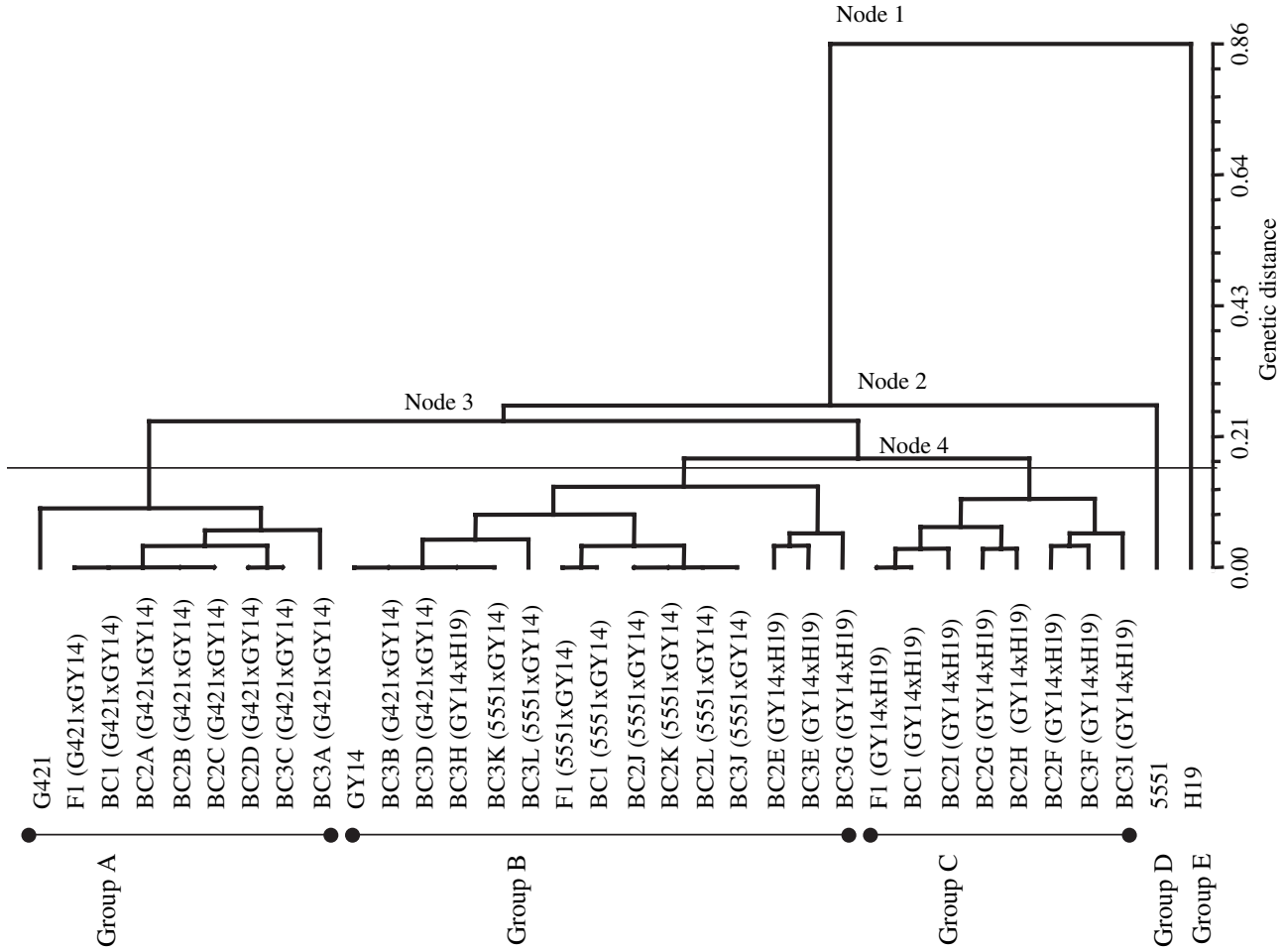

Fig. 3: Cluster analysis (by UPGMA) of cucumber (*Cucumis sativus* L.) lines (H-19, WI 5551, GY14, and G421) and their cross progeny grouped using genetic distances (Jaccard's coefficient) as estimated by 22 genetic markers (17 SSR and five SCAR)

population would be more effective for estimating rate of progress during backcrossing in these germplasm arrays for the establishment of ED threshold values. Such sampling procedures are often cost-prohibitive. Thus, the application of GD estimation in itself may not provide the descriptive power needed and/or be too expensive for PVP and ED application in this species.

For PVP, the description of the novel entity must be established *de novo* and its distinctness, uniformity, and stability unequivocally proved. Only molecular technologies that can assist in defining and/or understanding phenotypic expression as it relates to genotypic differences are of value to PVP and for establishing ED threshold values in a crop species. Sequence analysis may provide for such an opportunity in cucumber. In our study, an analysis of two distantly related lines (H-19 and G421; Fig. 1) provided unequivocal evidence for distinct nucleotide differences (Fig. 4). The allelic difference between these lines is defined by the SCAR marker $L18_{600}$ (Fazio 2001) and has potential legal import as it is associated with an economically important trait (i.e., multiple lateral branching; Fazio et al. 2003). This association is an example of and defines functional GD for this trait in cucumber. Functional GD is derived from marker-trait associations that estimate economic difference and is defined by the point at which the cumulative effect of marker-trait associations (e.g., quantitative trait loci) in a mapped genome accounts for a significant amount of the observed variation for a trait (Staub 1999). Such differences can be used in PVP if a specific and unique marker can be designed that is closely linked to the target trait. In this instance, a unique SNP marker (L18-1-H19B) was designed (i.e., both the sense and antisense primers) so that the 3′ end of the oligonucleotides matched only the SNP allele evident in H-19 (Fazio 2001). The difference between H-19 and G421 defined by SNP L18-1-H19B characterizes a functional GD because it is associated directly with high lateral branch number that is unique to H-19, i.e. possesses the allele for lateral branch number (Table 2; Fazio et al. 2003). Line H-19 was derived by self-pollination of line AR 79-75 (synom. 'Little John') released by the University of Arkansas, Fayetteville, in 1993, and characteristic of H-19. This character-specific DNA difference could have been used in support of the patent application of line AR 79-75 had it existed at the time of its release. A more robust support of patent application would be contributed if a marker were part of the functional gene itself.

In species having a narrow genetic base, such as cucumber, the level of genetic polymorphisms may restrict the use of molecular markers for PVP and the establishment of ED

```
Consensus ACCACCCACCATCATAGTCAAATAAGAAATGAAAGACAATAGAATAATAACTCACCTC
     H-19 ACCACCCACCATCATAGTCAAATAAGAAATGAAAGACAATAGAATAATAACTCACCTC
     G421 ACCACCCACCATCATAGTCAAATAAGAAATGAAAGACAATAGAATAATAACTCACCTC

Consensus AAGATTTGAGAAATTCCAATTGTACTGATAAACARAATCTAGTAGATCCCACTGATAA
     H-19 AAGATTTGAGAAATTCCAATTGTACTGATAAACAGAATCTAGTAGATCCCACTGATAA
     G421 AAGATTTGAGAAATTCCAATTGTACTGATAAACAAAATCTAGTAGATCCCACTGATAA

Consensus TATAAAACCAAWAAAGCGTTAGAGAMAAAATTCAATTGCTCACCCAGRAGAGAACAAA
     H-19 TATAAAACCAAAAAAGCGTTAGAGACAAAATTCAATTGCTCACCCAGGAGAGAACAAA
     G421 TATAAAACCAATTAAAGCGTTAGAGAAAAAATTCAATTGCTCACCCAGAAGAGAACAAA

Consensus -ACCCAACATTATAAATAAGACACAAACCTCTGTTCCCAYGGGGAATACCWCCTTCCA
     H-19 -ACCCAACATTATAAATAAGACACAAACCTCTGTTCCCATTGGGGAATACCACCTTCCA
     G421 GACCCAACATTATAAATAAGACACAAACCTCTGTTCCCACGGGGAATACCTCCTTCCA

Consensus CAAGTCTTCCTGAATAAGACAAGTAAGTAAACTCGAYCAGACATTATCCAGTACAAAT
     H-19 CAAGTCTTCCTGAATAAGACAAGTAAGTAAACTCGACCAGACATTATCCAGTACAAAT
     G421 CAAGTCTTCCTGAATAAGACAAGTAAGTAAACTCGATCAGACATTATCCAGTACAAAT

Consensus AACATCCCGAAGCACTAAAATYWA--TAGCGCACACATATCTACCACGTTTACAACGA
     H-19 AACATCCCGAAGCACTAAAATCAAAATAGCGCACACATATCTACCACGTTTACAACGA
     G421 AACATCCCGAAGCACTAAAATTTA--TAGCGCACACATATCTACCACGTTTACAACGA

Consensus GCAACRAACTCCTCAAAAACCAAAAACATTCAATTACTACAACAAATAAGCAAAATGA
     H-19 GCAACAAACTCCTCAAAAACCAAAAACATTCAATTACTACAACAAATAAGCAAAATGA
     G421 GCAACGAACTCCTCAAAAACCAAAAACATTCAATTACTACAACAAATAAGCAAAATGA

Consensus GGAAGCACATTCCAGAAAGTTGCTTGCTACACTACTGTTCTATAAGCTTTACAAAGTT
     H-19 GGAAGCACATTCCAGAAAGTTGCTTGCTACACTACTGTTCTATAAGCTTTACAAAGTT
     G421 GGAAGCACATTCCAGAAAGTTGCTTGCTACACTACTGTTCTATAAGCTTTACAAAGTT

Consensus ATGCTACTAAAACCMAAATAACGAAARGAAAACCCCTCGAGATTGGGATAAA-TGTAA
     H-19 ATGCTACTAAAACCAAAATAACGAAAAGAAAACCCCTCGAGATTGGGATAAA-TGTAA
     G421 ATGCTACTAAAACCCAAATAACGAAAGGAAAACCCCTCGAGATTGGGATAAAATGTAA

Consensus CGAATGCAGAATAAWAATAATAACCRAATTGCGCTTATCCTGGAAGTATTCCGGTACT
     H-19 CGAATGCAGAATAATTAATAATAACCAAATTGCGCTTATCCTGGAAGTATTCCGGTACT
     G421 CGAATGCAGAATAAAAATAATAACCGAATTGCGCTTATCCTGGAAGTATTCCGGTACT

Consensus GAAAGAGGTTTGSAAGCCTCGACTCGCTTGGCTCGGGTGGGTGG
     H-19 GAAAGAGGTTTGGAAGCCTCGACTCGCTTGGCTCGGGTGGGTGG
     G421 GAAAGAGGTTTGCAAGCCTCGACTCGCTTGGCTCGGGTGGGTGG
```

Fig. 4: Sequence comparison of amplified fragments of $L18_{600}$ sequence-characterized amplified region (SCAR) maker (Fazio 2001) using DNA from cucumber lines H-19 and G421 as templates. Polymorphic nucleotide sites are in bold and italic where Y = C or T, R = A or G, M = A or C, W = A or T, and S = C or G (R, M, W, and S are specific site designations where nucleotide differences reside)

threshold values. The use of mapped single nucleotide differences for genetic characterization will likely provide opportunities to define specific functional distances that have potential for PVP in cucumber and other species having well-characterized genetic maps. It is likely that without an expanded, genetically robust standard marker array (e.g., 50 codominant markers), ED threshold values in cucumber will be difficult to define, and thus SNPs will likely be required for the appraisal of genetic difference in this species.

References

ASSINSEL, 1999: http://www.worldseed.org/intellectual_property.htm.

Bernet, G. P., S. Bramardi, D. Calvache, E. A. Carbonell, and M. J. Asins, 2003: Applicability of molecular markers in the context of protection of new varieties of cucumber. Plant Breeding **122,** 146—152.

Dijkhuizen, A., W. C. Kennard, M. J. Havey, and J. E. Staub, 1996: RFLP variability and genetic relationships in cultivated cucumber. Euphytica **90,** 79—89.

Fazio, G., 2001: Comparative Study of Marker-Assisted and Phenotypic Selection and Genetic Analysis of Yield Components in Cucumber. PhD Thesis. University of Wisconsin-Madison, Madison, WI.

Fazio, G., J. E. Staub, and S. M. Chung, 2002: Development and characterization of PCR markers in cucumber (*Cucumis sativus* L.). J. Am. Soc. Hort. Sci. **127,** 545—557.

Fazio, G., J. E. Staub, and M. R. Stevens, 2003: Genetic mapping and QTL analysis of horticultural traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. Theor. Appl. Genet. **107,** 864—874.

Gilliland, T. J., R. Coll, E. Calsyn, M. DeLoose, M. J. T. van Eijk, and I. Roldan-Ruiz, 2000: Estimating genetic conformity between related ryegrass (Lolium) varieties. 1. Morphology and biochemical characterization. Mol. Breed. **6,** 569—580.

Horejsi, T., and J. E. Staub, 1999: Genetic variation in cucumber (*Cucumis sativus* L.) as assessed by random amplified polymorphic DNA. Genet. Res. Crop Evol. **46,** 337—350.

Horejsi, T., J. Box, and J. E. Staub, 1999: Efficiency of RAPD to SCAR marker conversion and their comparative PCR sensitivity in cucumber. J. Am. Soc. Hort. Sci. **124,** 128—135.

Hubalek, Z., 1982: Coefficients of association and similarity, based on binary (presence/absence) data: an evaluation. Biol. Rev. Camb. Philos. Soc. **57,** 669—689.

Jaccard, P., 1908: Nouvelles reserches sur la distribution florale. Bull. Soc. Vaud. Sci. Nat. **44,** 223—270.

Jackson, D. A., K. M. Somers, and H. H. Harvey, 1989: Similarity coefficients: measurements of co-occurrence and association or simply measures of occurrence? Am. Nat. **133,** 436—453.

Janowitz, M. F., 1980: Similarity measures on binary data. Syst. Zool. **29,** 342—359.

Knerr, L. D., J. E. Staub, D. J. Holder, and B. P. May, 1989: Genetic diversity in *Cucumis sativus* L. assessed by variation at 18 allozyme coding loci. Theor. Appl. Genet. **78,** 119—128.

Kruskal, J. B., and M. Wish, 1978: Multidimensional scaling. In: Sage University Paper Series on Quantitative Applications in the Social Sciences. Sage Publ., Beverly Hills, CA.

Lombard, V., P. Dubreuil, C. Dillmann, and C. Baril, 2001: Genetic distance estimators based on molecular data for plant registration and protection: A review. In: C. Baril, C. Dore, and F. Dosba (eds), Proceedings of the International Symposium on Molecular Markers for Characterizing Genotypes and Identifying Cultivars in Horticulture. Acta Hortic. **546,** 55—63.

Maniatis, T., E. F. Fritsch, and J. Sambrook, 1982: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Meglic, V., and J. E. Staub, 1996: Genetic diversity in cucumber (*Cucumis sativus* L.): II. An evaluation of selected cultivars released between 1846 and 1978. Genet. Res. Crop Evol. **46,** 547—558.

Meglic, V., F. Serquen, and J. E. Staub, 1996: Genetic diversity in cucumber (*Cucumis sativus* L.): I. A reevaluation of the US germplasm collection. Genet. Res. Crop Evol. **46,** 533—546.

Mliki, A., J. E. Staub, Z. Sun, and A. Ghorbel, 2003: Genetic diversity in African cucumber (*Cucumis sativus* L.) provides potential for germplasm enhancement. Genet. Res. Crop Evol. **50,** 461—468.

Nei, M., 1973: Analysis of gene diversity in subdivided populations. Proc. Natl Acad. Sci. USA **70,** 3321—3323.

de Riek, J., 2001: Are molecular markers strengthening plant variety registration and protection? In: J. nan Huylenbroeck, and E. van Bokstaele (eds), Proc. XX Eucarpia Symposium on New Ornamentals. Acta Hortic. **552,** 215—233.

Rohlf, F. J., 1997: NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System. Exeter Software, New York.

Serquen, F. C., J. Bacher, and J. E. Staub, 1997: Mapping and QTL analysis of a narrow cross in cucumber (*Cucumis sativus* L.) using random amplified polymorphic DNA markers. Mol. Breed. **3,** 257—268.

Smith, J. S. C., D. S. Ertl, and B. A. Orman, 1995: Identification of maize varieties. In: C. W. Wrigley (ed.), Identification of Food-Grain Varieties, 252—264. American Association of Cereal Chemists, St Paul, MN.

Sorensen, T., 1948: A method for establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analysis of the vegetation on Danish commons. Biologiske Skrifter **5,** 1—34.

Spooner, D. M., J. Tivang, J. Nienhuis, J. T. Miller, D. S. Douches, and M. Contreras, 1996: Comparison of four molecular markers in measuring relationships among wild potato relatives *Solanum* section *Etuberosum* (subgenus *Potato*). Theor. Appl. Genet. **92,** 532—540.

Staub, J. E., 1999: Intellectual property rights, genetic markers, and hybrid seed production. J. New Seeds **1,** 39—65.

Staub, J. E., and J. Bacher, 1997: Cucumber as a processed vegetable. In: D. S. Smith, J. N. Cash, W. K. Nip, and Y. H. Hui (eds), Processing Vegetables: Science and Technology IV, 129—193. Technomic Publishing Co., Inc., Lancaster, PA.

Staub, J. E., and V. Meglic, 1993: Molecular genetic markers and their legal relevance for cultigen discrimination: a case study in cucumber. HortTechnology **3,** 291—300.

Staub, J. E., A. Gabert, and T. C. Wehner, 1996a: Plant variety protection: a consideration of genetic distance. HortScience **31,** 1086—1091.

Staub J. E., J. Bacher, and J. D. Poetter, 1996b: Sources of potential errors in the application of random amplified polymorphic DNA in cucumber. HortScience **31,** 262–266.

Staub, J. E., F. C. Serquen, and J. D. McCreight, 1997: Genetic diversity in cucumber (*Cucumis sativus* L.): III. An evaluation of Indian germplasm. Genet. Res. Crop Evol. **44,** 315—326.

Staub, J. E., F. Dane, K. Reitsma, G. Fazio, and A. Lopez-Sees, 2002: The formation of test arrays and a core collection in (*Cucumis sativus* L.) using phenotypic and molecular marker data. J. Am. Soc. Hort. Sci. **127,** 558—567.

Steel, R. G. D., and J. H. Torrie, 1980: Principles and Procedures for Statistics. A Biometrical Approach. McGraw-Hill Book Company, New York.

UPOV, 1990: Guidelines for the Conduct of Tests for Distinctness, Homogeneity, and Stability. Document UPOV TG/4/7. UPOV, Geneva, Switzerland.

UPOV, 1991: International Convention for the Protection of New Varieties of Plants. UPOV Publication 221(E), 19 March 1991. UPOV, Geneva, Switzerland.

Widrlechner, M. P., L. D. Knerr, J. E. Staub, and K. R. Reitsma, 1992: Biochemical evaluation of germplasm regeneration methods for cucumber, *Cucumis sativus* L. FAO/IBPGR Plant Genet. Resources Newslett. **88/89,** 1—4.